npg

# ORIGINAL ARTICLE

# Physical activity classification utilizing SenseWear activity monitor in manual wheelchair users with spinal cord injury

SV Hiremath[1,2], D Ding[1,2,3], J Farringdon[4], N Vyas[4] and RA Cooper[1,2,3]

**Study design:** Validation.

**Objectives:** The primary aim of this study was to develop and evaluate activity classification algorithms for a multisensor-based SenseWear (SW) activity monitor that can recognize wheelchair-related activities performed by manual wheelchair users (MWUs) with spinal cord injury (SCI). The secondary aim was to evaluate how the accuracy in activity classification affects the estimation of energy expenditure (EE) in MWUs with SCI.

**Setting:** University-based laboratory.

**Methods:** Forty-five MWUs with SCI wore a SW on their upper arm and participated in resting, wheelchair propulsion, arm-ergometry and deskwork activities. The investigators annotated the start and end of each activity trial while the SW collected multisensor data and a portable metabolic cart collected criterion EE. Three methods including linear discriminant analysis, quadratic discriminant analysis (QDA), and Naïve Bayes (NB) were used to develop classification algorithms for four activities based on the training data set from 36 subjects.

**Results:** The classification accuracy was 96.3% for QDA and 94.8% for NB when the classification algorithms were tested on the validation data set from nine subjects. The average EE estimation errors using the activity-specific EE prediction model were $5.3 \pm 21.5\%$ and $4.6 \pm 22.8\%$ when the QDA and NB classification algorithms were applied, respectively, as opposed to $4.9 \pm 20.7\%$ when 100% classification accuracy was assumed.

**Conclusion:** The high classification accuracy and low EE estimation errors suggest that the SW can be used by researchers and clinicians to classify and estimate the EE for the four activities tested in this study among MWUs with SCI.

Spinal Cord (2013) 51, 705–709; doi:10.1038/sc.2013.39; published online 21 May 2013

## INTRODUCTION

Activity classification using wearable activity monitors among the ambulatory population has been well documented.[1–5] The benefits of detecting physical activities (PAs) using wearable devices include the ability to track regular PA, provide accurate energy expenditure (EE) estimation and assist in behavioral modifications that may lead to a healthier active lifestyle in community settings.[6–10] However, there are only a limited number of studies that have detected and classified PAs performed by individuals who rely on wheelchairs for mobility using wearable devices.[11–13] Identification of wheelchair-related PAs using wearable devices provide not only all the benefits mentioned above but also pertinent information on the functional use of upper limbs, an important factor of upper limb pain and injury prevalent in wheelchair users.[14] The clinical practice guideline 'Preservation of Upper Limb Function Following Spinal Cord Injury,' published by the Paralyzed Veterans of America, has indicated that minimizing the frequency of upper extremity use in wheelchair users during repetitive tasks such as wheelchair propulsion can decrease the risk factor for repetitive strain injury and/or wrist pain.[14]

Previous research by Postma et al.[11] showed that a wearable activity monitor consisting of six accelerometers and two electrocardiogram electrodes connected to a portable data recorder (0.7 kg) was able to detect wheelchair propulsion in ten manual wheelchair users (MWUs) with spinal cord injury (SCI). The results showed that wheelchair propulsion episodes were detected with an overall agreement, sensitivity and specificity of 92%, 87% and 92%, respectively. In another study, French et al.[13] showed that wheelchair propulsion patterns, surface types and self-propulsion versus external pushing of a wheelchair could be detected using two dual-axis accelerometer-based eWatches secured to the wrist and the wheelchair's frame. The results in three persons without disabilities showed that the classification accuracy rates varied from 80 to 90% for arcing versus non-arcing propulsion patterns, carpet versus tile surfaces and self-propulsion versus external pushing using classification algorithms such as k-nearest neighbor and support vector machines. Along similar lines, Ding et al.[12] studied activity classification in 27 MWUs performing a series of representative activities of daily living in a semi-structured setting with an eWatch and a wheel rotation

[1]Department of Veterans Affairs, Human Engineering Research Laboratories, VA Pittsburgh Healthcare System, Pittsburgh, PA, USA; [2]Department of Rehabilitation Science and Technology, University of Pittsburgh, Pittsburgh, PA, USA; [3]Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA, USA and [4]BodyMedia Inc., Pittsburgh, PA, USA

Correspondence: Dr D Ding, Department of Veterans Affairs, Human Engineering Research Laboratories, VA Pittsburgh Healthcare System, 6425 Penn Avenue, Suite 400, Pittsburgh, PA 15206, USA.

E-mail: dad5@pitt.edu

datalogger placed on the wrist and the wheelchair's wheel, respectively. The results indicated that k-nearest neighbor, support vector machine, Naïve Bayes (NB) and decision tree (C4.5) classification algorithms could classify the activities into self-propulsion, external pushing and sedentary activity with an accuracy of 89.4–91.9%. The studies discussed here focused specifically on detecting propulsion activity versus other activities with activity monitoring systems composed of multiple components.

The primary objective of this study was to develop and evaluate machine learning-based classification algorithms to detect PAs including resting, wheelchair propulsion, arm ergometer exercises and deskwork performed by MWUs with SCI based on data collected from an off-the-shelf multisensor-based sensewear (SW) activity monitor. Our previous research has shown that an activity-specific EE prediction model consisting of four EE estimation equations for the four types of PAs mentioned above had smaller EE estimation errors than a general model consisting of only one EE estimation equation applied for all the activities.[15] Therefore, in order to use the activity-specific EE prediction model in the field, we first need to detect the four types of PAs. Our secondary aim was to evaluate how the activity classification accuracy affects the performance of the activity-specific EE prediction model for MWUs with SCI described in our previous work.[15]

## MATERIALS AND METHODS

### Experimental protocol
The study was approved by the institutional review board at the University of Pittsburgh and the VA Pittsburgh Healthcare System. The target population of this study was MWUs with SCI. Participants were recruited through the institutional review board approved registries, flyers and advertisements in print media. Convenience sampling was used to recruit participants who expressed an interest in the study. Little or no research has been published on validating activity monitors for EE estimation among MWUs with SCI. Power analysis using a correlational design with $\alpha = 0.05$ (two-tail) and medium effect size ($r = 0.4$) indicated that a total of 40 participants will provide a statistical power of 74%.[16] On the basis of this estimation, in this study we recruited 45 MWUs with SCI to take part in the study and provide a written informed consent before their participation in the study. The data collection for the study took place between February 2009 and May 2011. Participants were included if they were between 18 and 60 years of age, used a manual wheelchair as a primary means of mobility, had an SCI, were at least 6 months post-injury and were able to use an arm-ergometer for exercise. Participants were excluded if they were unable to tolerate sitting for 4 h, had active pelvic or thigh wounds or failed to obtain their primary care physician's consent to participate in the study. The study required the participants to pay one visit to the Human Engineering Research Laboratories, University of Pittsburgh to complete the data collection. All 45 participants who provided written informed consent participated in the study.

The research study protocol has been described in detail elsewhere.[15,17] As part of the pre-activity session, the participants answered a demographics questionnaire and had their heights and weights measured. During the activity session, the participants took part in resting and three other activities including wheelchair propulsion, arm-ergometer exercises and deskwork. The three activities were counterbalanced and the trials within each activity were randomized to counter order effects. During the activity session, all participants wore a SW activity monitor on their right upper arm over the triceps and a Cosmed K4b2 portable metabolic cart (COSMED srl, Rome, Italy). The participants performed each activity trial for a maximum period of 8 min, with a resting period of 5–10 min between activity trials and a period of 30–40 min between activities. During the wheelchair propulsion activity, the participants propelled their wheelchairs for two trials of 2 and 3 mph on a stationary dynamometer, and a trial of 3 mph on a flat-tiled surface. The arm-ergometer exercises included two trials at 60 r.p.m. with 20 and 40 W of resistance and a trial at 90 r.p.m. with 40 W of resistance. During the deskwork

session, the subjects typed on a computer for 4 min and read a book for another 4 min.

### Instrumentation and data collection
The SW activity monitor was used to collect the average, the mean absolute difference (variability of upper limb motion) and the number of peaks (turning points of upper limb) in transverse and longitudinal accelerations sampled at 32 Hz and recorded at 16 Hz; and the average galvanic skin response (skin conductance due to moisture or sweat), skin temperature and near body temperatures sampled at 32 Hz and recorded at 1 min. The multisensor data from the SW was retrieved using the InnerView Research software 7.0 (Bodymedia Inc., Pittsburgh, PA, USA). In addition, a portable K4b2 metabolic cart was synchronized with the SW and used to collect the criterion EE. The EE in terms of kcal min$^{-1}$ was retrieved using the Cosmed K4b2 software (version 9.0). The investigators annotated the start and end of each activity trial during data collection, which was further used as the reference for developing and testing of the classification algorithms.

### Data analysis
The first step of developing an activity classification algorithm was to separate the data into a training data set and a validation data set. A stratified approach with the injury level (paraplegia versus tetraplegia) as the stratified variable was used to select 80% of the participants into the training data set and 20% into the validation data set. The total amount of activity time was 1645 min (about 27.4 h) including 1319 min (about 22.0 h) in the training data set ($n = 36$) and 326 min (about 5.4 h) in the validation data set ($n = 9$).

The next step was to extract a set of features, which are statistical measures, used to distinguish between the four types of activities. The feature data included characteristic information such as the mean, the mean absolute difference and the number of peaks per minute that were directly obtained from various sensors in the SW activity monitor. In addition, linear and nonlinear features using the multisensor data from SW were calculated on the basis of statistical characteristics, such as time domain features, biomechanical and physiological features specific to PAs.[15] We chose a 1-min window size (duration or period) for feature estimation to be consistent with the EE estimation. The features obtained from the SW and the estimated features resulted in a feature space of thousands of variables for the PA classification. We also manually labeled each 1-min activity segment as belonging to one of the four categories, that is, wheelchair propulsion, arm ergometry, resting and deskwork based on the annotations, which served as a reference for training and testing the activity classification algorithms. The data collected from the SW was processed through data analysis programs written in MATLAB (The Mathworks, Inc., Natick, MA, USA).

We then developed three activity classification algorithms based on the training data set using machine learning algorithms including linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and NB. For each classification algorithm, we performed the leave-one-subject-out (LOSO) and sixfold by-subject cross-validation to select the most appropriate features and evaluate the classification algorithm's performance. The LOSO cross-validation method leaves one subject out and then develops the model on the remaining subjects. The model developed on these remaining subjects is evaluated by the left-out subject. This procedure was repeated 36 times, as there were 36 subjects in training group. The sixfold by-subject cross-validation method is similar to LOSO, except that the subjects are split into six random groups (or folds), and each time a group is left out and the models are developed on the remaining five groups. The sixfold cross-validation was repeated six times as the total participants in the training data set were 36. In addition to cross-validation, the performance of the three activity classification algorithms was also evaluated using the validation data set. Several performance measures were calculated including per-minute precision (true positive/(true positive + false positive)), recall (true positive/(true positive + false negative)), specificity (true negative/(true negative + false positive)) and overall accuracy ((true positive + true negative)/(number of the cases)).[18] Precision indicates the proportion with which the detected activity is correct. Recall, also known as sensitivity, is the proportion of actual activities that are correctly identified. Specificity is the proportion of activities not performed that are correctly

identified, or in other words it is the classification algorithm's ability to distinguish actual true-negative cases. Overall accuracy is the overall performance of the algorithm. We also evaluated how the performance of the activity classification algorithms affected the EE estimation using the activity-specific EE prediction model that was previously developed.[15] In our previous work, the EE estimation based on the activity-specific prediction model assumed 100% activity classification accuracy. However, in this study we evaluated the performance of the activity-specific EE prediction model based on the actual classification results. Similar to our previous work,[15] the estimated EE was compared with the criterion EE from the metabolic cart by calculating the minute-by-minute mean absolute error and the mean-signed error.

## RESULTS

The participants included 37 males and 8 females with a m.s.d. age of 40.2 (11.0) years, weight of 78.5 (21.9) kg, height of 178.2 (8.6) cm and manual wheelchair usage of 13.8 (9.1) years. Thirty-eight participants had paraplegia (SCI of T4 and below) and seven participants had tetraplegia (SCI of T3 and above). Detailed demographics has been discussed in our previous work.[15] Table 1 shows the performance of the LDA, QDA and NB classification algorithms applied to the training data set ($n = 36$) using the LOSO and sixfold by-subject cross-validation methods. The results showed that the classification accuracy was less dependent on the algorithms, but more dependent on the type and number of features. For the sake of brevity, we have chosen to present detailed results of the QDA and NB

classification algorithms. Table 2 shows the classification performance in terms of the precision, recall, sensitivity and overall accuracy of the QDA and NB classification algorithms using four features in the validation data set. The overall classification performance was 96.3% and 94.8% for QDA and NB classification algorithms, respectively. Table 3 shows the confusion matrix, which is a visual representation of the actual or true activity and the activity detected by the classification algorithm. The results from Table 3 indicate that the misclassification often occurred between wheelchair propulsion and arm ergometry exercises, which involve repetitive upper extremity usage. Furthermore, Table 4 shows the EE estimation errors including the mean absolute error and mean-signed error for the validation data set ($n = 9$) when the activity-specific EE prediction model was used in conjunction with the QDA or NB classification algorithms with four features.

## DISCUSSION

Accessible activity monitors in wheelchair users will allow users themselves, researchers and clinicians to track regular PA, EE estimation, PA levels in community settings and functional use of upper limbs, which is related to pain and injury prevalence in wheelchair users. Results from this study indicate that the SW activity monitor along with custom machine learning classification algorithms, such as LDA, QDA and NB can be used to classify wheelchair-related PAs in MWUs. Compared with the study conducted by Postma et al.[11] where six activity monitors were used to detect wheelchair propulsion episodes from a series of activities, we used a single SW activity monitor to achieve a higher classification accuracy (96% for QDA classification algorithm versus 92%) with a larger number of subjects ($n = 45$ versus $n = 10$). Similarly, the classification algorithms discussed here outperformed those in the previous studies by Ding et al.[12] and French et al.[13], who classified wheelchair-related PAs by using two devices in smaller number ($n = 27$) of wheelchair users and three non-wheelchair users, respectively.

Several strategies were used to reduce overfitting during the classification algorithm development. As shown in Table 1, the classification accuracy improved with an increased number of features, indicating that a reasonable number of features are necessary to classify multiple PAs. Given the number of participants in the study, we chose to use a small feature set including four features for further analysis of the classification algorithms, as we wanted to strike a balance between accuracy and overfitting of the classification algorithms to unseen participants. Furthermore, the results showed that the LOSO cross-validation technique that tends to have higher variance and lower bias in a small sample had similar performance to

**Table 1 Classification performance in terms of the overall accuracy (%) for the LDA, QDA and NB classification algorithms to detect four wheelchair-related activities with varied number of features using the LOSO and sixfold by-subject cross-validation methods in the training data set**

| Cross-validation | Machine learning algorithms | Number of features | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 10 |
| LOSO | LDA | 70.28 | 82.13 | 88.64 | 91.56 | 92.88 | 94.67 |
| | QDA | 74.08 | 83.77 | 90.69 | 93.44 | 94.39 | 96.20 |
| | NB | 74.08 | 82.76 | 91.56 | 93.67 | 93.95 | 93.95 |
| Sixfold by-subject | LDA | 69.95 | 78.75 | 84.85 | 88.84 | 91.78 | 94.39 |
| | QDA | 73.44 | 82.51 | 91.12 | 93.38 | 94.40 | 94.66 |
| | NB | 73.44 | 82.78 | 91.99 | 93.54 | 94.26 | 95.47 |

Abbreviations: LDA, linear discriminant analysis; LOSO, leave-one-subject-out; NB, Naïve Bayes; QDA, quadratic discriminant analysis.

**Table 2 Classification performance in terms of the precision (true-positive rate), recall (sensitivity), specificity (true-negative rate) and overall accuracy (%) of the QDA and NB classification algorithms using four features to detect the four wheelchair-related activities in the validation data set**

| Class | % QDA | | | | % NB | | | |
|---|---|---|---|---|---|---|---|---|
| | Precision | Recall | Specificity | Overall | Precision | Recall | Specificity | Overall |
| Resting | 100.0 | 97.1 | 100.0 | 99.7 | 100.0 | 97.1 | 100.0 | 99.7 |
| Propulsion | 92.8 | 98.3 | 95.7 | 96.6 | 94.1 | 94.1 | 96.6 | 95.7 |
| Arm ergometry | 99.3 | 93.7 | 99.5 | 96.9 | 95.0 | 93.7 | 96.2 | 95.1 |
| Deskwork | 94.1 | 100.0 | 99.3 | 99.4 | 91.4 | 100.0 | 99.0 | 99.1 |

Abbreviations: NB, Naïve Bayes; QDA, quadratic discriminant analysis.

**Table 3 Confusion matrix for the QDA and NB classification algorithms using four features to classify the four wheelchair-related activities in the validation data set**

| Class | QDA | | | | NB | | | |
|---|---|---|---|---|---|---|---|---|
| | Resting | Propulsion | Arm ergometry | Deskwork | Resting | Propulsion | Arm ergometry | Deskwork |
| Resting | 33 | 0 | 0 | 1 | 33 | 0 | 0 | 1 |
| Propulsion | 0 | 116 | 1 | 1 | 0 | 111 | 7 | 1 |
| Arm ergometry | 0 | 9 | 133 | 0 | 0 | 7 | 133 | 2 |
| Deskwork | 0 | 0 | 0 | 32 | 0 | 0 | 0 | 32 |

Abbreviations: NB, Naïve Bayes; QDA, quadratic discriminant analysis.

**Table 4 EE estimation error in terms of the mean absolute error and mean-signed error for the validation data set when the activity-specific EE prediction model was used in conjunction with the QDA or NB classification algorithms with four features**

| Class | QDA | | | | NB | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean absolute error per minute | | Mean-signed error per minute (s.d.) | | Mean absolute error per minute | | Mean-signed error per minute (s.d.) | |
| | kcal | % | kcal | % | kcal | % | kcal | % |
| Resting | 0.2 | 18.2 | 0.0 (0.3) | −4.3 (21.4) | 0.2 | 18.2 | 0.0 (0.3) | −4.3 (21.4) |
| Propulsion | 0.6 | 16.5 | 0.2 (0.7) | 1.9 (22.4) | 0.7 | 18.7 | 0.0 (1.0) | −0.1 (25.3) |
| Arm ergometry | 0.9 | 18.8 | 0.7 (1.0) | 11.6 (20.1) | 0.9 | 18.8 | 0.7 (1.0) | 11.8 (20.1) |
| Deskwork | 0.2 | 13.4 | 0.0 (0.2) | −0.4 (16.1) | 0.2 | 13.4 | 0.0 (0.2) | −0.4 (16.1) |
| All activities | 0.7 | 17.4 | 0.4 (0.9) | 5.3 (21.5) | 0.7 | 18.2 | 0.3 (1.0) | 4.6 (22.8) |

Abbreviations: EE, energy expenditure; NB, Naïve Bayes; QDA, quadratic discriminant analysis.

the sixfold by-subject cross-validation technique. This led us to use the LOSO cross-validation for classification algorithm development, which helps improve the generalizability of the classification algorithms to unseen participants. The four features for the QDA classification algorithm were: the resultant acceleration, and three other variables derived from the mean absolute difference and number of peaks of the transverse acceleration. Similarly, the four features for the NB classification algorithm were: the resultant acceleration and three other variables derived from the mean absolute difference of the transverse acceleration, and mean absolute difference and number of peaks of the longitudinal acceleration. The features chosen by both the QDA and NB classification algorithms included directional, total motion and frequency of upper arm movement information from the SW's accelerometer, indicating that the classification algorithms were sensitive to movement-based variables when classifying the wheelchair-related PAs. Even though the QDA classification algorithm yielded slightly higher accuracy than NB, the NB classification algorithm is computationally simpler and has greater potential for real-time activity classification.

In our previous work,[15] we developed an activity-specific EE prediction model, which involves detecting the type of PA before applying a specific EE estimation equation for the detected PA. However, the previous work evaluated the model performance assuming the types of PAs that can be detected and classified with 100% accuracy. With over 95% classification accuracies yielded by the QDA and NB classification algorithms, we found that the performance of the activity-specific EE prediction model was minimally affected by the actual classification results. The previous study showed that the mean absolute error and mean-signed error for all activities were 16.8% and 4.9 ± 20.7%, respectively.[15] In this study, the mean absolute error and mean-signed error for all activities were

17.4% and 5.3 ± 21.5% for the QDA classification algorithm, respectively, and 18.2% and 4.6 ± 22.8% for the NB classification algorithms, respectively. The results in Table 3 also showed that the wheelchair propulsion and arm ergometry activities were occasionally misclassified by QDA and NB classification algorithms; yet the misclassification may not significantly affect the EE prediction as the two activities have similar EE. Further, the activity-specific EE estimation equations and the classification algorithms share some common variables including the mean absolute difference of transverse acceleration, and the mean absolute difference and average number of peaks of longitudinal acceleration.[15]

One limitation of this study is the small number of PAs tested in the protocol. In addition, the activities were performed in a controlled laboratory setting and prescribed in a precise manner such as propelling a wheelchair and exercising with an arm ergometer at a certain speed and/or intensity. Future studies should evaluate a larger number of PAs in the home and community of MWUs. To our knowledge, there is no device that can be directly used by wheelchair users to classify PAs and estimate EE. We chose to investigate the potential of SW activity monitor in this population owing to its ready availability in the market and multisensor capabilities.

## CONCLUSION
Availability of physical activity monitors for MWUs can empower them to monitor everyday PA participation and EE, and make informed decisions toward healthier behaviors. The high classification accuracy of the QDA and NB classification algorithms and the low EE estimation errors when using the actual classification results suggest that the SW activity monitor can be used to classify and estimate the EE for the four activities tested in this study among MWUs with SCI.

1   Mannini A, Sabatini AM. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors* 2010; **10**: 1154–1175.

2   Vyas N, Farringdon J, Andre D, Stivoric J. *Machine learning and sensor fusion for estimating continuous energy expenditure, in the 2011 Conference on Innovative Applications in Artificial Intelligence*. AAAI press: San Francisco, CA, 2011.

3   Andre D, Wolf DL. Recent advances in free-living physical activity monitoring: a review. *J Diabetes Sci Technol* 2007; **1**: 760–767.

4   Albinali F, Intille S, Haskell W, Rosenberger M. Using wearable activity type detection to improve physical activity energy expenditure estimation. In: *Proceedings of the 12th Conference on Ubiquitous Computing*. ACM Press: New York, 2010, pp 311–320.

5   Gyllensten IC, Bonomi AG. Identifying types of physical activity with a single accelerometer: evaluating laboratory trained algorithms in daily life. *IEEE Trans Biomed Eng* 2011; **58**: 2656–2663.

6   Berntsen S, Hageberg R, Aandstad A, Mowinckel P, Anderssen SA, Carlsen KH et al. Validity of physical activity monitors in adults participating in free living activities. *Br J Sports Med* 2010; **44**: 657–664.

7   Johannsen DL, Calabro MA, Stewart J, Frankie W, Rood JC, Welk GJ. Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Med Sci Sports Exerc* 2010; **42**: 2134–2140.

8   Polzien KM, Jakicic JM, Tate DF, Otto AD. The efficacy of a technology-based system in a short-term behavioral weight loss intervention. *Obesity* 2007; **15**: 825–830.

9   Jakicic JM, Marcus M, Gallagher K, Randall C. Evaluation of the sensewear pro armband to assess energy expenditure during exercise. *Med Sci Sports Exerc* 2004; **36**: 897–904.

10  Sui X, Meriweather RA, Hand GA, Wilcox S, Dowda M, Blair SN. Electronic feedback in a diet and physical activity-based lifestyle intervention for weight loss: randomized controlled trial. *AHA 50th Annual EPI/NPAM Joint Conference* 2010.

11  Postma K, Berg-Emons van den HJG, Bussmann JBJ, Sluis TAR, Bergen MP, Stam HJ. Validity of the detection of wheelchair propulsion as measured with an activity monitor in patients with spinal cord injury. *Spinal Cord* 2005; **43**: 550–557.

12  Ding D, Hiremath S, Chung Y, Cooper R. Detection of wheelchair user activities using wearable sensors. In: Stephanidis C (ed.). *Universal Access in Human-Computer Interaction. Context Diversity. Lecture Notes in Computer Science*, Vol. 6767. Springer: Berlin, Heidelberg, 2011: 145–152.

13  French B, Tyamagundlu D, Siewiorek DP, Smailagic A, Ding D. Towards a Virtual Coach for manual wheelchair users. *Proceedings of the 12th IEEE International Symposium on Wearable Computers* 2008, pp 77–80.

14  Paralyzed Veterans of America Consortium for Spinal Cord Medicine. Preservation of upper limb function following spinal cord injury: A clinical practice guideline for health-care professionals. *J Spinal Cord Med* 2005; **28**: 434–470.

15  Hiremath SV, Ding D, Farringdon J, Cooper RA. Predicting energy expenditure of manual wheelchair users with spinal cord injury using a multi-sensor based activity monitor. *Arch Phys Med Rehab* 2012; **93**: 1937–1943.

16  Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*, (ed 2) Prentice-Hall: Upper Saddle River, NJ, 2000.

17  Hiremath SV, Ding D. Evaluation of activity monitors in manual wheelchair users with paraplegia. *J Spinal Cord Med* 2011; **34**: 110–117.

18  Kohavi R, Provost F. Glossary of terms. *Mach Learn* 1998; **30**: 271–274.